When Independent Gaussian Models Break Down: Characterizing Regime-Dependent Modeling Failures in ϕ^4 Theory

Anish Bhat*
University of California, San Diego
a2bhat@ucsd.edu

Ryo Ide*
University of California, San Diego
ride@ucsd.edu

Zihan Zhao
University of California, San Diego
ziz078@ucsd.edu

Abstract

In practical physical systems, modeling assumptions of Gaussianity and basis independence break down due to self-interactions. We study a specific instance of one-dimensional ϕ^4 theory on a lattice, analyzing how the interaction strength and system size jointly affect the marginal and joint distributions of frequency-based representation of the field (i.e., Fourier modes). We find that models relying on Gaussian and independent Fourier modes fail primarily due to structured dependencies rather than marginal non-Gaussianity, since individual modes become approximately Gaussian despite mode coupling growing with size. Based on this, we identify three distinct regimes that delineate where traditional methods remain effective and where more expressive models are needed. Our results provide a computationally simple diagnostic to establish when Gaussian models are insufficient, and derive structural requirements that motivate future nonlinear model designs.

1 Introduction

Methods that assume an independent basis for modeling physical systems decompose complex systems into simple components, making them a crucial tool for machine learning modeling of physical phenomena. However, this relies on the assumption that the physical system can be decomposed into modes that act independently, which is often violated in real systems.

One such system is the one-dimensional quartic interaction (1D ϕ^4 Theory), a type of interaction in quantum field theory. Self-interactions are deliberately introduced in the modeling action principle by adding $\lambda\phi^4$ where ϕ is the field and λ is the coupling constant that determines interaction strength.

While certain bases, such as a Fourier basis, diagonalize non-interacting field theories and imply statistical independence between modes, interactions in ϕ^4 theory can break this assumption through mode coupling. Currently, it is unknown to what extent Fourier-based methods can still accurately describe ϕ^4, at what scales, and when independence-based Gaussian representations become insufficient.

Our main contributions are: (1) testing the limits of current methods based on independent representations; (2) testing independence-agnostic data-driven methods; (3) identifying the coupling regimes where methods fail; (4) identifying properties that future nonlinear models may need to express.

*Equal contribution.

Conference on Physics and AI at Stanford University (PAI 2026)

2 Problem setup

To address these questions, we compare Fourier-based Gaussian models with data-driven baselines using samples from a Markov Chain-Monte Carlo (MCMC) simulation of ϕ^4 theory, which is well established in this subject [2, 4, 6]. Our full experimental setup is available at `https://anonymous.4open.science/r/phi4-PAI-2026-Stanford-F12B/README.md`.

Field Distribution. We consider a 1D lattice scalar field $\phi \in \mathbb{R}^N$ with periodic boundary conditions. Our MCMC implementation targets the following Boltzmann distribution $p(\phi) \propto e^{-S[\phi]}$, where the action is $S[\phi] = \sum_{i=1}^{N} \left[\frac{1}{2}(\phi_{i+1} - \phi_i)^2 + \frac{1}{2}m^2\phi_i^2 + \lambda\phi_i^4\right]$ [7]. $S[\phi]$ encodes the energy cost of a field configuration; the Boltzmann distribution weights configurations by how energetically favorable they are, with λ controlling the strength of self-interaction.

Note that when $\lambda = 0$, the distribution $p(\phi)$ can be fully factorized into a pure Gaussian. In general, increasing λ increases interaction strength.

Sampling. Samples are generated using the Metropolis-Hastings MCMC algorithm on a field with coupling constants $\lambda \in \{0.1, 1.0, 5.0, 10.0, 20.0, 50.0, 100.0\}$.

Observables. The power spectrum is computed via the discrete Fourier transform $\tilde{\phi}_k = \mathcal{F}[\phi]_k$, where $P(k) = \mathbb{E}[|\tilde{\phi}_k|^2]$ [12].

3 Models and baselines

Fourier Model. As a baseline, we use a Fourier-based method traditionally used to model free theories without self-interactions, which we apply here to ϕ^4 Theory. We assume that Fourier modes are independent Gaussian variables $\tilde{\phi}_k \sim \mathcal{N}(0, \sigma_k^2)$, with σ_k^2 estimated from data. Samples are generated by independently sampling each mode and applying the inverse Fourier transform.

Principal Component Analysis Model. As a data-driven baseline, we use PCA to find $z = U^\top(\phi - \bar{\phi})$, where U are the eigenvectors of $\Sigma = U\Lambda U^T$[1].

Full-Gaussian Model. As an independence-agnostic baseline, we fit $\tilde{\phi} \sim N(0, \Sigma)$ directly to data, testing how far second-order Gaussian behavior persists as interaction increases.

Masked Autoregressive Flow. For a nonlinear baseline, we train a masked autoregressive flow (MAF) on Fourier-space samples [13], learning a nonlinear transformation from the base Gaussian in order to represent arbitrary joint structure including 4th-order inter-mode dependencies.

4 Experimental design

Normalized Coupling Metric First, we define spectral energy $E_k = |\tilde{\phi}_k|^2$. E_k is a random variable over the sample ensemble, so $\mathrm{Cov}(E_k, E_{k'})$ is estimated across samples. Second, computing covariance $\Sigma^{(E)} = \mathrm{Cov}(E_k, E_{k'})$, we define the normalized coupling metric as $C = \frac{\|\Sigma^{(E)} - \mathrm{diag}(\Sigma^{(E)})\|_F}{\|\Sigma^{(E)}\|_F}$. For jointly Gaussian variables, statistical independence implies zero pairwise covariance (i.e., $C = 0$). C therefore measures how much the spectral energy variables violate independence, with $C = 1$ indicating that off-diagonal structure dominates. Since $E_k = |\tilde{\phi}_k|^2$ is quadratic in $\tilde{\phi}_k$, the covariance $\mathrm{Cov}(E_k, E_{k'})$ involves fourth-order moments of the field, making C a fourth-order statistic.

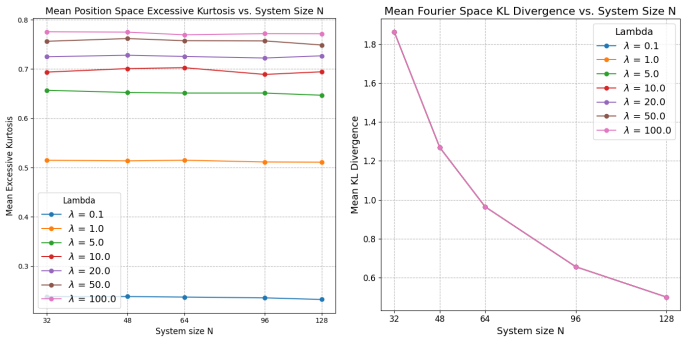
Spectral Error Metric To measure model performance, we define normalized spectral error $\epsilon = \frac{\sum_k |\hat{P}(k) - P(k)|}{\sum_k P(k)}$, where $P(k)$ is the empirical power spectrum and $\hat{P}(k)$ is the predicted power spectrum from generated samples.

Excess Kurtosis. To quantify deviations from Gaussianity in marginal distributions, we compute the excess kurtosis $K(x)$ for position-space variables ϕ_i and Fourier modes $\tilde{\phi}_k$ [11]. A negative (positive) K indicates lighter (heavier) tails, providing a simple diagnostic of marginal non-Gaussianity.

Kullback-Leibler Divergence. To measure deviations from Gaussian structure, we compute the Kullback–Leibler (KL) Divergence $D_{\mathrm{KL}}(p\|q)$ between the empirical distribution $p(x)$ and a Gaussian reference $q(x)$ with matched mean and covariance [14]. We estimate this quantity on 1D marginals (per-site and per-mode), quantifying marginal non-Gaussianity more sensitively than kurtosis. Unlike joint metrics such as the coupling metric C, $D_{\mathrm{KL}}(p\|q)$ does not capture mode coupling, allowing us to explicitly separate marginal effects from inter-mode dependencies.

5 Results

Gaussianity Analysis. First, we examine marginal Gaussianity in position space via excess kurtosis K_{position}. At weak coupling, near-Gaussian behavior is observed (see Figure 1a). K_{position} is largely invariant with system size N, indicating that local statistics are determined primarily by interaction strength. While local marginals become strongly non-Gaussian at large λ, they are effectively fixed once the system is sufficiently large. This suggests that Gaussianity-agnostic density models (e.g., Variational Autoencoders (VAEs) or Normalizing Flows) may better model local field distributions by capturing this non-Gaussian structure [5, 9, 10].



(a) Mean K_{position} Magnitude vs. N (b) Mean $D_{\text{KL, Fourier}}$ vs. N

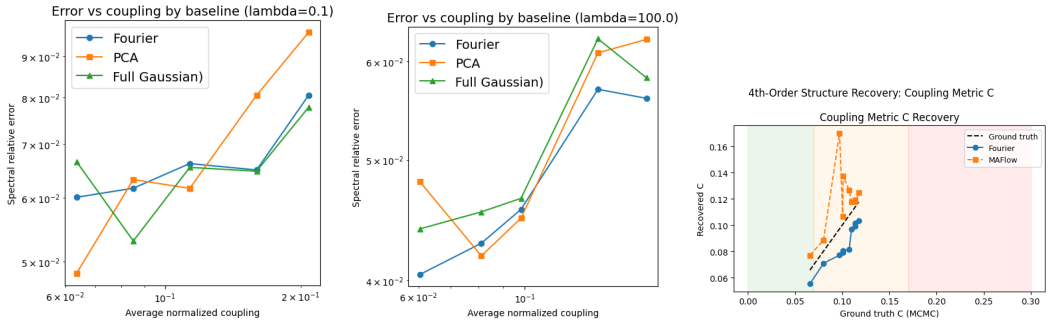
Figure 1: 1a: Empirically, for $\lambda = 0.1$, $K_{\text{position}} \approx -0.23$, indicating near-Gaussian behavior. Largely independent of N, it saturates at -0.77, near the theoretical limit of -0.81 given by standard ϕ^4 theory. 1b: Empirically, $D_{\text{KL, Fourier}}$ gradually decreases to $D_{\text{KL, Fourier}} \approx 0.4$. Similarly, K_{Fourier} approaches ≈ 0.14 with N, likewise indicating near-Gaussian marginals. Note that curves overlap.

In Fourier space, both excess kurtosis K_{Fourier} and KL divergence $D_{\text{KL, Fourier}}$ decrease with N and approach stable values with nearly identical behavior across all λ (see Figure 1b). This indicates that individual Fourier modes become approximately Gaussian in distribution. While this trend is consistent with Central-Limit-Theorem-type behavior arising from averaging over many degrees of freedom, the convergence occurs even as interactions remain present. Thus, marginal Gaussianity alone does not imply statistical independence between modes. Despite the increasingly Gaussian Fourier marginals, we observe that the spectral error ϵ grows monotonically with coupling C (see Figure 2a–2b). Since performance is evaluated across Gaussian-family models here, this result suggests that accurate marginal statistics are insufficient to guarantee accurate joint modeling. Instead, performance appears increasingly limited by the presence of structured inter-mode dependencies as coupling grows. This motivates future investigation of models representing higher-order joint structure, such as VAEs and Normalizing Flows [3, 8].

Regime-Dependent Behavior. We identify three empirical regimes based on the measured coupling value C (see Figure 2a–2b): (i) a weak-coupling regime ($C \lesssim 0.07$) where spectral error remains small and independent representations perform well; (ii) an intermediate-coupling regime ($0.07 \lesssim C \lesssim 0.17$) where increasing coupling is accompanied by a corresponding increase in spectral error; and (iii) a strong-coupling regime ($C \gtrsim 0.17$) where both coupling and spectral error saturate. In this

regime, $C \in (0.17, 0.18)$ and $\epsilon \in (0.05, 0.06)$ for most tested settings. Although spectral error is a second-order observable, its monotonic relationship with the fourth-order coupling metric C suggests that it remains sensitive to the same interaction-driven structure captured by C. These observations suggest that C separates regimes more directly than λ or N alone.

To examine whether recovering the joint 4th-order structure is possible, we evaluate the MAF on C (see Figure 2c). While the Fourier baseline systematically underestimates C (0.088 vs. ground truth 0.105 in the intermediate regime), the MAF consistently overestimates C (0.128), outperforming Fourier in only half of intermediate-regime settings.



(a) $\lambda = 0.1$ (b) $\lambda = 100.0$ (c) MAF vs. Fourier

Figure 2: 2a-2b Spectral relative error vs. normalized coupling across all baselines. Similar trends are observed for spectral error with respect to system size N. This is consistent with coupling increasing with N: across all λ, C ranges from $C(N = 32) \approx 0.06$ to $C(N = 128) \approx 0.2$. 2c Recovered coupling C vs. ground truth C across all baselines. The Fourier baseline systematically underestimates C by construction; the MAF recovers higher coupling values but overshoots, outperforming Fourier in half of intermediate-regime settings.

6 Limitations

Restricted Physical Setup. We study a 1D lattice with periodic boundary conditions and a fixed bare mass m^2. Higher dimensions, alternative boundary conditions, and the symmetry-breaking phase transition of ϕ^4 as a function of m^2 may alter the regime structure identified here. A joint sweep over (m^2, λ) is left for future work.

Diagnostic and Model Scope. The metric C is a 4th-order statistic and does not constrain higher cumulants (≥ 6) that may matter at strong coupling. Our MAF baseline probes 4th-order structure but exhibits training instability, so we cannot distinguish whether residual strong-coupling error is fundamental or capacity-limited. We do not estimate the MCMC sampling noise floor at fixed (λ, N), so we cannot rule out that the observed saturation of ϵ is partially a noise artifact.

7 Conclusion

We identify three coupling regimes in 1D ϕ^4 theory and find a separation between marginal simplicity and joint complexity: Fourier modes become roughly Gaussian marginally as N grows, while joint coupling C grows by a factor of $\sim$3. Gaussian baselines fit marginals well by construction, suggesting that residual error increasingly reflects unexpressed joint higher-order structure.

Our MAF results show that nonlinear models can capture inter-mode coupling that Gaussian baselines miss by design, though instability in training limits consistent recovery. Our results suggest that successful models require that they (i) preserve approximate marginal Gaussianity of Fourier modes, and (ii) capture 4th-order joint cumulants between modes. These requirements explain why fully factorized Gaussian families, full-covariance Gaussian models, and linear bases leave residual error in this regime, which further motivates coupling-layer normalizing flows and nonlinear latent-variable models. While demonstrated in 1D ϕ^4 theory, the coupling diagnostic C is expected to arise in any system where tunable self-interactions compete with a natural independent-mode decomposition.

References

[1] Hervé Abdi and Lynne J. Williams. Principal component analysis. *WIREs Computational Statistics*, 2(4):433–459, 2010. doi: 10.1002/wics.101.

[2] M. S. Albergo, G. Kanwar, and P. E. Shanahan. Flow-based generative models for Markov chain Monte Carlo in lattice field theory. *Phys. Rev. D*, 100:034515, 2019. doi: 10.1103/PhysRevD.100.034515.

[3] M. S. Albergo, D. Boyda, D. C. Hackett, G. Kanwar, K. Cranmer, S. Racanière, D. J. Rezende, and P. E. Shanahan. Introduction to normalizing flows for lattice field theory. *arXiv preprint*, 2101.08176, 2021. URL `https://arxiv.org/abs/2101.08176`.

[4] Richard C. Brower and Pablo Tamayo. Embedded dynamics for ϕ^4 theory. *Phys. Rev. Lett.*, 62:1087–1090, Mar 1989. doi: 10.1103/PhysRevLett.62.1087. URL `https://link.aps.org/doi/10.1103/PhysRevLett.62.1087`.

[5] Michele Caselle, Elia Cellini, Alessandro Nada, and Marco Panero. Stochastic normalizing flows for lattice field theory. *PoS*, LATTICE2022:005, 2023. doi: 10.22323/1.430.0005.

[6] M Hasenbusch. A Monte Carlo study of leading order scaling corrections of ϕ^4 theory on a three-dimensional lattice. *Journal of Physics A: Mathematical and General*, 32(26):4851–4865, January 1999. ISSN 1361-6447. doi: 10.1088/0305-4470/32/26/304. URL `http://dx.doi.org/10.1088/0305-4470/32/26/304`.

[7] Max Jansen and Kilian Nickel. ϕ^4 theory on the lattice, March 2011. Lecture notes, University of Bonn. URL `https://www.hiskp.uni-bonn.de/uploads/media/phi4.pdf`.

[8] G. Kanwar, M. S. Albergo, D. Boyda, K. Cranmer, D. C. Hackett, S. Racanière, D. J. Rezende, and P. E. Shanahan. Equivariant flow-based sampling for lattice gauge theory. *Phys. Rev. Lett.*, 125:121601, 2020. doi: 10.1103/PhysRevLett.125.121601.

[9] Diederik P. Kingma and Max Welling. An introduction to variational autoencoders. *Foundations and Trends in Machine Learning*, 12(4):307–392, 2019. doi: 10.1561/2200000056. URL `https://doi.org/10.1561/2200000056`.

[10] Ivan Kobyzev, Simon J.D. Prince, and Marcus A. Brubaker. Normalizing flows: An introduction and review of current methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(11):3964–3979, November 2021. ISSN 1939-3539. doi: 10.1109/tpami.2020.2992934. URL `http://dx.doi.org/10.1109/TPAMI.2020.2992934`.

[11] NIST/SEMATECH. Measures of skewness and kurtosis. NIST/SEMATECH e-Handbook of Statistical Methods, Section 1.3.5.11, 2012. URL `https://www.itl.nist.gov/div898/handbook/eda/section3/eda35b.htm`.

[12] Alan V. Oppenheim and Ronald W. Schafer. *Discrete-Time Signal Processing*. Pearson, Upper Saddle River, NJ, 3rd edition, 2010.

[13] George Papamakarios, Theo Pavlakou, and Iain Murray. Masked autoregressive flow for density estimation. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 2335–2344. Curran Associates, Inc., 2017. URL `https://arxiv.org/abs/1705.07057`.

[14] Cosma Shalizi. Shannon entropy and Kullback-Leibler divergence. Lecture notes, Carnegie Mellon University, 36-754, 2006. URL `https://www.stat.cmu.edu/~cshalizi/754/2006/notes/lecture-28.pdf`.